A NOVEL STATISTICAL BASED APPROACH TO NON-LINEAR MODEL UPDATING USING RESPONSE FEATURES

John F. Schultze, François M. Hemez, Scott W. Doebling, and Thomas A. Butler

Engineering Analysis Group (ESA-EA), M/S P946
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

ABSTRACT

This research presents a new method to improve analytical model fidelity for non-linear systems. The approach investigates several mechanisms to assist the analyst in updating an analytical model based on experimental data and statistical analysis of parameter effects. The first is a new approach at data reduction called ***feature extraction***. This is an expansion of the 'classic' update metrics to include specific phenomena or characters of the response that are critical to model application. This is an extension of the familiar linear updating paradigm of utilizing the eigen-parameters or frequency response functions (FRFs) to include such devices as peak acceleration, time of arrival or standard deviation of model error. The next expansion of the updating process is the inclusion of statistical based parameter analysis to quantify the effects of uncertain or significant effect parameters in the construction of a ***meta-model***. This provides indicators of the statistical variation associated with parameters as well as confidence intervals on the coefficients of the resulting meta-model. Also included in this method is the investigation of linear parameter effect screening using a partial factorial variable array for simulation. This is intended to aid the analyst in eliminating from the investigation the parameters that do not have a significant variation effect on the feature metric. Finally an investigation of the model to replicate the measured response variation is examined.

1. MOTIVATION

The updating and validation of complex non-linear models to reflect not only 'real world' data but also its variability is of strong interest in the aerospace, automotive and aviation industries ([1], [2], [3]). The higher objective is to improve confidence in the model within and beyond the experimental range, since it is often impractical to test over the full operational range of a system. An additional objective is to develop an understanding and identification of the relation between significant input parameters, such as Young's Modulus, and critical response data components (features). Construction of a 'meta-model' between the input and output is developed to aide in this understanding and to reduce the computational load of investigating parameter variation.

2. METHODOLOGY

A procedure is developed to determine dominant input variables, construct and evaluate the meta-model, including mechanisms to judiciously select input variable levels, and investigate the relation between analytical predictions and experimental results for not only error but also statistical distribution properties.

Current model updating methods in structural dynamics are generally based on linear assumptions and do not have a quantifiable confidence index of model components. Several methods use either the measured eigen-parameters or FRFs. These techniques commonly attempt to either map the experimental information to the model space or the converse. This results in the confounding of system information through data expansion or condensation. Identified modeling errors are associated with specific parameters or physical regions of a model, often without a critical analytic justification. There is normally little evaluation, from either a Design of Experiments (DoE) [4] or statistical approach to quantify the model update mechanism for its range of applications and confidence intervals.

Development of a new method based on use of response 'features' and a DoE approach parameter variation analysis in the updating of analytical models is examined. This method is applicable to time-varying non-linear systems where classical methods often do not succeed. This method also provides for confidence indications of model components.

A 'feature' is an identified physically and/or analytically significant quantity derived from the response data. This could be as simple as the peak level of a single response record or a more coupled metric such as the standard deviation of model error over the entire response space. The former is one of the metric evaluated in this paper and the latter is currently under investigation for a different model. A 'feature', by its nature, is a general term and is specified by the analyst. Under this guideline the traditional update choices of eigen-parameters would qualify as features, though their application is only meaningful for linear systems. Selection of features that are amenable to the initial linear screening and have justifiable ***physical*** meaning is often one of the greater challenges with this method.

3. APPLICATION

A flowchart of the proposed method is shown in Figure 1. The development presented will follow this guide. The method is iterative in nature and the selection of features is dependent on the analyst's goals and insights. In some instances, iterations will be performed within steps and in most cases at least some redefinition/refinement of parameters and their levels, or input settings, is necessary.

This technique has been applied to systems of varying complexity at Los Alamos National Laboratory (LANL)[6]. The example illustrated in this presentation will be the drop-test model and experiments for characterizing visco-elastic material behavior.

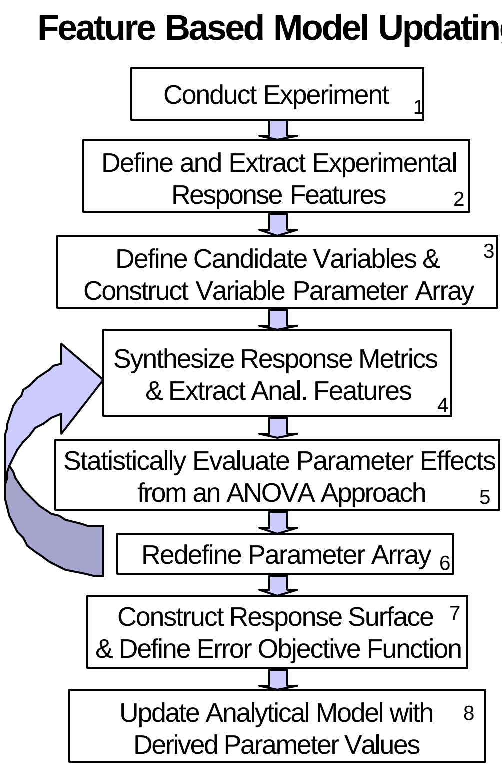

Figure 1. Model Updating Method Overview

The following sections step though this process.

4. IMPLEMENTATION

The first step is to conduct initial experimentation for the identification of response features and selection of preliminary input variables. An input variable is chosen for several possible reasons, first if it is known to vary in operation and experimentation and expected to have a significant impact on the response feature, it is included. Second if it is not readily measurable, but expected to have impact, it may be considered. The experimentalist and the analyst determine other candidates.

Figure 2 shows the measured input and output for one set of the experimental tests and illustrates the variability in 'identical' experiments. It is important to recognize that the variability in the experimentation must be taken into consideration not only in determining the proper adjustments to the model, but also as a limiting criteria for the update procedure itself. It should be recognized that it is impractical to refine a model such that it has lower variability than a test of the same design point.

Also illustrated in the figure below is the concept of a response feature. In this case the peak level of acceleration and its time of arrival are the selected response features. These were selected with input from material scientists as critical for material characterization.

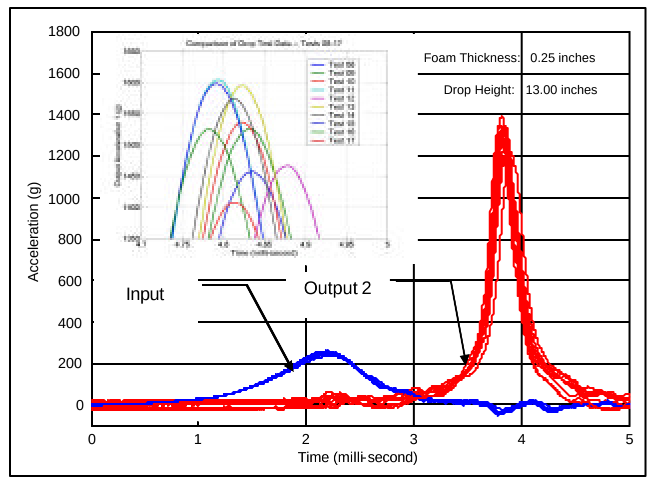

Figure 2. Experimental Variation, Input and Output 2

In this experiment, there are eight (8) input variables, resulting in an eight-dimensional sample space. It is impractical to construct a full factorial experiment for this high of dimensionality. Therefore we use the sampling method of orthogonal arrays. This mechanism allows for linear effects to not be confounded (aliased) with each other and provides and overall efficient sampling of the design space. In Figure 3, the variability of the Abaqus [5] Finite Element Analysis (FEA) model is shown for the 81-run Orthogonal Array ensemble. It is clear that the original model spans the range of experimental response, which is a crucial indicator of it usefulness. If the model and the chosen input variable ranges do not span the experimental set that will be used in updating, serious questions exist about the appropriateness of the model and the selected input levels. Re-examining the input ranges, and expanding as appropriate, with the restriction that the ranges not extend into non-realizable values, can often remedy this situation.

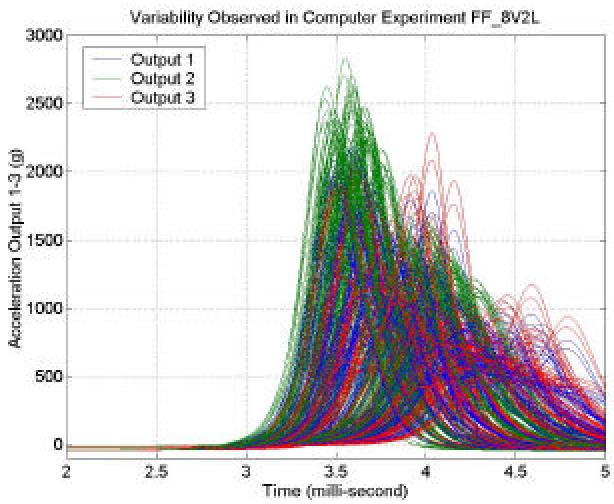

Figure 3. Model Results for the OA81 Run Ensemble

One of the benefits of constructing a meta-model of the system and the initial linear screening is identification of the statistically significant input parameters. Often the initial set of input variables is inclusive, which prevent the analysis of multiple designs when the simulation code is resource expensive. Through the DoE design of the input array [7], [8] and subsequent analysis of variation (ANOVA) of the interrelation between input levels and output features the parameter set is reduced. This reduced set is then used to create a meta-model or response surface of the system. Often additional computer simulations of the model are performed to provide adequate input levels for higher order polynomial models in the ANOVA.

Table 1. Dominant Parameters

	OA27	OA27r	OA81	OA81r	FF256	FF256r
Peak G Loc 1	C,H	A, C, B, F	C, A, F, B, H	C, A, B, F	A, C, H, B, E, G	C, A, B, F, G
TOA 1	C, A, B, F, G, H	C, F, A, E	C, A, G, F	C, E, A, F	C, F, A, G, E, B	C, A, F, E, B
Peak G Loc 2	C, H	C, F, B, G	C, F, H,	C, F, A, G E, H	C, H, F, E, G, B	C, F, B, G, E
TOA 2	C, F, G, A, B	C, E, F, G, B, A	C, G, F, E, A	C, E, F, G, B, A	C, F, G, E	C, E, F, G, A, B
Peak G, Loc 3	H	C, A, B	C, A, B, F, H	C, A, B, F, H	C, B, A, H, E	C, A, B, F, G
TOA 3	C, B, F	C, B, F, E	C, B, F, H	C, B, F, E	C, B, F, G, H	C, B, F, E

Table 1 shows the results of the three initial simulation run-sets. The letters A-H represent chosen input variables and their order represents there corresponding significant linear participation. For example in run-set OA81, the analysis shows that parameters C, A, F, B and H have an significant impact in decreasing magnitude.

If scores are then assigned to each parameter and its order in a model, an evaluation of contributing input variables can be made. In this case, 4 points were assigned for 1st place, 3 for 2nd and so on. The results are presented in Table 2. Clearly D has no effect, H was shown to be a false positive and G is significantly lower than the remaining variables, so A, B C, E and F were retained for further analysis and model construction.

Table 2. Parameter Contribution Scores, Sorted by Dataset

	OA27	OA27r	OA81	OA81r	FF256	FF2 56r	Total
A	4	9	9	10	8	9	49
B	5	9	6	7	7	9	43
C	20	23	24	24	23	24	138
D	0	0	0	0	0	0	0
E	0	5	1	7	2	5	20
F	6	11	11	10	10	11	59
G	2	2	5	2	4	2	17
H	10	0	3	0	5	0	18

A meta-model was formed for each output feature for several simulation runs. To keep the number of simulations at a tractable number two approaches were used, either the simulation had five (5) variables and three (3) levels, (243 runs), or four (4) variables and four (4) levels, (256 runs).

Figure 4 shows the results of the simulation runs. Both models span the experimental range and perform generally well. It is clear though from this figure and Figure 5 that the four variable model has less error.

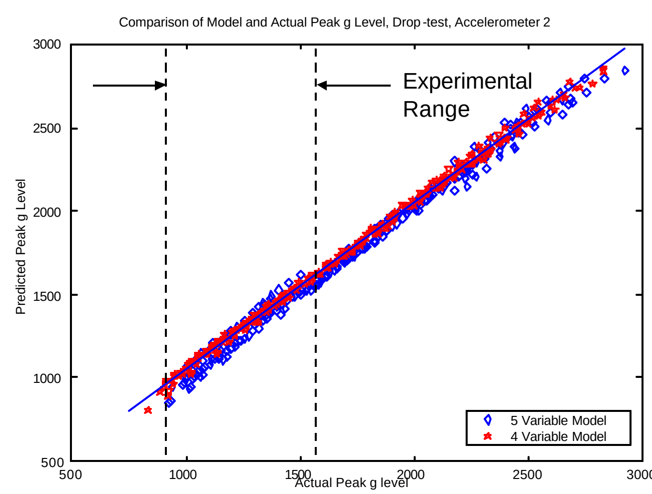

Figure 4. Model Comparison for Peak Acceleration Prediction

In this case the model had A, B, C and F retained. While information theory states that we should see an error reduction with smaller model dimension for essentially the same number of data points (runs in this case) it was shown in other research not presented here that these four variables are the best sub-set.

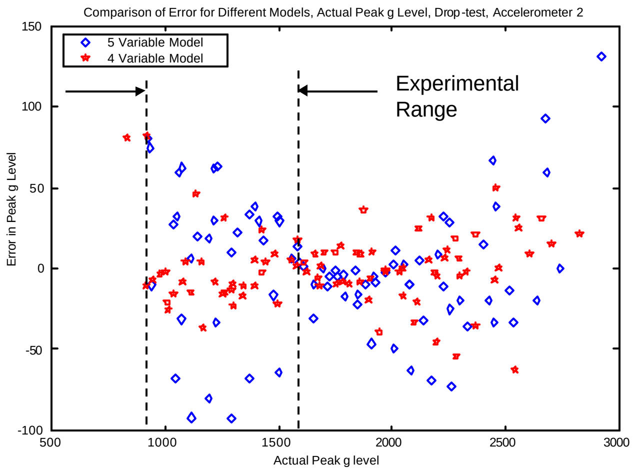

Figure 5. Error in Model Prediction of Peak G, Loc. 2

5. APPLICATION OF META-MODELS

The next step is to infer from test data the optimal values of the input parameters. The procedure followed when the investigation is restricted to four parameters: (the two angles of impact (A, B), the bolt pre-load (C) and the input scaling (F)) will now be introduced.

Since a smaller number of input parameters are retained (4 out of 8), a localized computer experiment can be designed to provide a better resolution in the area of interest. The area of interest is here defined as the region in the multi-dimensional feature space where responses measured during testing are located. As mentioned previously, a full factorial matrix designed from a Taguchi array formed of four levels for each input parameter is analyzed. Then, fast running models are fit to the data. Equation (1) illustrates one of the models typically obtained for the peak acceleration response at sensor #2:

$$\ddot{x}_2^{\text{peak}} = \{-1{,}538.2 \quad 43.6 \quad 288.4 \quad 2.4 \quad 2{,}552.8 \quad -391.3 \quad \cdots \quad -307.1 \quad 665.7 \quad -0.5 \quad -452.4 \quad 1.5\} * \{1 \quad ?_1 \quad ?_2 \quad P_{\text{bolt}} \quad a_I \quad ?_1^2 \quad \cdots \quad ?_2^2 \quad ?_1 * ?_2 \quad ?_2 * P_{\text{bolt}} \quad ?_2 * a_I \quad P_{\text{bolt}} * a_I\}^T$$

Equation 1

Instead of applying direct least-squares fitting to multi-dimensional polynomials, statistical based models are preferred because in addition to yielding computationally efficient meta-models, they also provide confidence intervals that can be used for assessing the model's goodness-of-fit. For example, each coefficient of the polynomial shown in Equation (1) is associated to statistics that show how dominant the corresponding effect is. Therefore, Equation (1) defines a family of models that can be re-sampled to account for omitted sources of uncertainty (round-off errors, environmental variability, etc.). Re-sampling the model essentially means that decisions would be based on properties of ensembles rather than a single model [9]. In addition, statistical models can be refined to optimize the statistical significance of each individual effect contribution, which may be more important than maximizing the overall goodness-of-fit to the data.

6. OBJECTIVE FUNCTION BASED OPTIMIZATION

To quantify the application of the model to explain the experimental response variation, an optimization was performed to adjust the analytical model to 'best-match' the physical test results. The objective function chosen was the distance squared between the experimental measured feature and the analytical prediction based on the meta-model using variables A, B, C, and F. Explicitly the general form of objective function was,

$$J(\Delta_A, \Delta_B, \Delta_C, \Delta_F) = \left(ExpFeature - AnalFeature(\Delta_A, \Delta_B, \Delta_C, \Delta_F)\right)^2$$

Equation 2

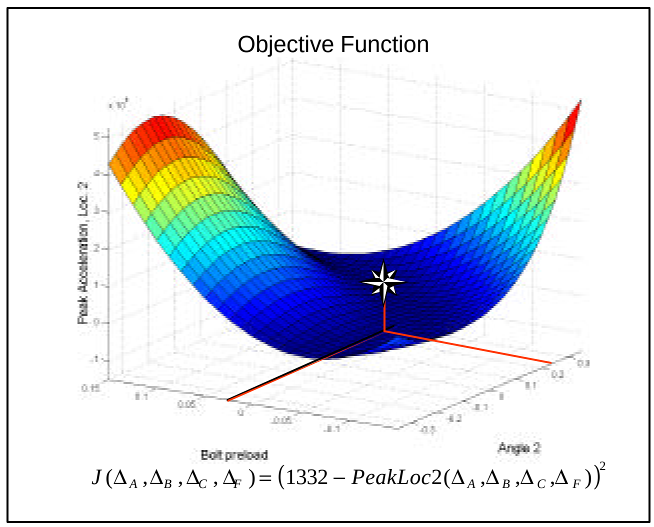

Figure 6. Surface of Objective Function for Optimization

Figure 6 illustrates, in normalized coordinates, a 2D response surface obtained from Equation (2). The mean acceleration response obtained from the data collected at sensor #2 is shown as a star. A straightforward optimization provides the optimal values of the input parameters. In this case, a pre-load equal to 200 psi (1.38 x 10^6 N/m^2) is obtained together with an impact angle equal to 0.7 degrees. Note that such an approach provides an optimized model capable of reproducing the mean response obtained from test data. It does not guarantee that the variance or other higher statistical moments are captured.

7. METHOD EVALUATION

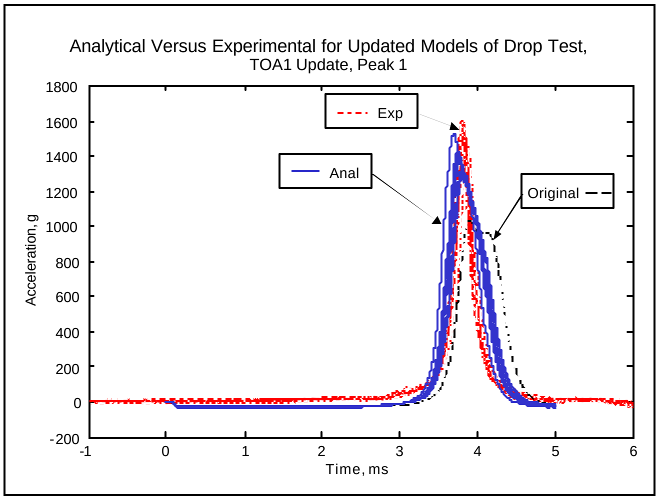

Figure 7. Comparison Between Analytical and Experimental Response Variation

Figure 7 confirms that the proposed procedure works well for this application. The experimental and corresponding 'meta-model optimized' analytical response has approximately the same variation and character. This validates the final analytical model and the update procedure well.

8. CONCLUSIONS

This research presents a new method to improve analytical model fidelity for non-linear systems. The approach investigates several mechanisms to assist the analyst in updating an analytical model based on experimental data and statistical analysis of parameter effects. The first is a new approach at data reduction called ***feature extraction***. The next expansion is the inclusion of statistical based parameter analysis to quantify the effects of uncertain or significant effect parameters in the construction of a ***meta-model***. The results from the linear screening, model refinement, variable variation and the response synthesis all are very promising. This should greatly aid the analyst in the update of large scale and non-linear models over the present methods.

9. REFERENCES

[1] Beardsley, P., Hemez, F.M., and Doebling, S.W., 'Updating Nonlinear Finite Element Models in the Time Domain,' *2nd International Workshop on Structural Health Monitoring,* Stanford University, Stanford, CA, Sep. 8-10, 1999, pp. 774-783.

[2] Hemez, F.M., Doebling, S.W., 'Validation of Nonlinear Modeling From Impact Test Data Using Probability Integration,' *18th IMAC, International Modal Analysis Conference*, San Antonio, TX, Feb. 7-10, 2000, pp. 124-132.

[3] Hemez, F.M., Doebling, S.W., 'Inversion of Structural Dynamics Simulations: State-of-the-art and Orientations of the Research,' *25th ISMA, International Conference in Noise and Vibration Engineering*, Leuven, Belgium, Sep. 13-15, 2000, pp. 403-413.

[4] Phadke, M. S., 'Quality Engineering using Robust Design', Prentice Hall, Englewood Cliffs, NJ, 1989

[5] **Abaqus/Explicit**, User's Manual, Version 5.8, Hibbitt, Karlsson & Sorensen, Pawtucket, RI, 1998.

[6] Butler, T. A., F. M. Hemez, J. F. Schultze, H. Soon, S. W. Doebling, 'Model Validation for a Complex Jointed Structure', *19th IMAC, International Modal Analysis Conference*, Kissimmee, FL, Feb. 5-8, 2001.

[7] Hedayat, A.S., Sloane, N.J.A., Stufken, J., **Orthogonal Arrays: Theory and Applications**, Springer-Verlag, NY, 1999.

[8] **Design-Expert® Software**, User's Guide, Version 6.0.1, Stat-Ease, Inc., MN, 2000.

[9] Rutherford, B., 'A Re-sampling-based Approach to Optimal Experimental Design for Computer Analysis of a Complex System,' *3rd DOE/MICS Workshop on the Predictability of Complex Phenomena*, Los Alamos, NM, Dec. 6-8, 1999.

[10] **Matlab© Version 5.3.1,** Mathworks, Cambridge, MA, 1999.